RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 09/508,658 D

Source: IFW016

Date Processed by STIC: 10/05/2005

ENTERED

IFW16

```
 3 <110> APPLICANT: KROHN, Kai
 4        HEINO, Maarit
 5        PETERSON, Part
 6        SCOTT, Hamish
 7        ANTONARAKIS, Stylianos
 8        LALIOTI, Maria D.
 9        SHIMIZU, Nobuyoshi D.
10        KUDOH, Jun D.
12 <120> TITLE OF INVENTION: NOVEL GENE DEFECTIVE IN APECED AND ITS USE
14 <130> FILE REFERENCE: u 012653-9
16 <140> CURRENT APPLICATION NUMBER: 09/508,658D
17 <141> CURRENT FILING DATE: 2000-11-03
19 <160> NUMBER OF SEQ ID NOS: 41
21 <170> SOFTWARE: PatentIn version 3.2
23 <210> SEQ ID NO: 1
24 <211> LENGTH: 2036
25 <212> TYPE: DNA
26 <213> ORGANISM: HOMO SAPIENS
29 <220> FEATURE:
30 <221> NAME/KEY: CDS
31 <222> LOCATION: (137)..(1771)
32 <223> OTHER INFORMATION: /product="AIR-1"
35 <220> FEATURE:
```

W--> ```
36 <221> NAME/KEY: mat peptide
```

```
37 <222> LOCATION: (1)..(545)
38 <223> OTHER INFORMATION: /product="AIR-1"
41 <400> SEQUENCE: 1
42 agaccgggga gacgggcggg cgcacagccg gcgcggaggc cccacagccc cgccgggacc    60
44 cgaggccaag cgaggggctg ccagtgtccc gggacccacc gcgtccgccc cagccccggg   120
46 tccccgcgcc caccccc atg gcg acg gac gcg gcg cta cgc cgg ctt ctg agg   172
47                    Met Ala Thr Asp Ala Ala Leu Arg Arg Leu Leu Arg
48                     1               5                  10
50 ctg cac cgc acg gag atc gcg gtg gcc gtg gac agc gcc ttc cca ctg      220
51 Leu His Arg Thr Glu Ile Ala Val Ala Val Asp Ser Ala Phe Pro Leu
52         15                  20                  25
54 ctg cac gcg ctg gct gac cac gac gtg gtc ccc gag gac aag ttt cag      268
55 Leu His Ala Leu Ala Asp His Asp Val Val Pro Glu Asp Lys Phe Gln
56     30                  35                  40
58 gag acg ctt cat ctg aag gaa aag gag ggc tgc ccc cag gcc ttc cac      316
59 Glu Thr Leu His Leu Lys Glu Lys Glu Gly Cys Pro Gln Ala Phe His
60 45                  50                  55                  60
62 gcc ctc ctg tcc tgg ctg ctg acc cag gac tcc aca gcc atc ctg gac      364
63 Ala Leu Leu Ser Trp Leu Leu Thr Gln Asp Ser Thr Ala Ile Leu Asp
```

```
64                  65                  70                  75
66 ttc tgg agg gtg ctg ttc aag gac tac aac ctg gag cgc tat ggc cgg        412
67 Phe Trp Arg Val Leu Phe Lys Asp Tyr Asn Leu Glu Arg Tyr Gly Arg
68                  80                  85                  90
70 ctg cag ccc atc ctg gac agc ttc ccc aaa gat gtg gac ctc agc cag        460
71 Leu Gln Pro Ile Leu Asp Ser Phe Pro Lys Asp Val Asp Leu Ser Gln
72         95                  100                 105
74 ccc cgg aag ggg agg aag ccc ccg gcc gtc ccc aag gct ttg gta ccg        508
75 Pro Arg Lys Gly Arg Lys Pro Pro Ala Val Pro Lys Ala Leu Val Pro
76      110                 115                 120
78 cca ccc aga ctc ccc acc aag agg aag gcc tca gaa gag gct cga gct        556
79 Pro Pro Arg Leu Pro Thr Lys Arg Lys Ala Ser Glu Glu Ala Arg Ala
80 125                 130                 135                 140
82 gcc gcg cca gca gcc ctg act cca agg ggc acc gcc agc cca ggc tct        604
83 Ala Ala Pro Ala Ala Leu Thr Pro Arg Gly Thr Ala Ser Pro Gly Ser
84                  145                 150                 155
86 caa ctg aag gcc aag ccc ccc aag aag ccg gag agc agc gca gag cag        652
87 Gln Leu Lys Ala Lys Pro Pro Lys Lys Pro Glu Ser Ser Ala Glu Gln
88                  160                 165                 170
90 cag cgc ctt cca ctc ggg aac ggg att cag acc atg tca gct tca gtc        700
91 Gln Arg Leu Pro Leu Gly Asn Gly Ile Gln Thr Met Ser Ala Ser Val
92         175                 180                 185
94 cag aga gct gtg gcc atg tcc tcc ggg gac gtc ccg gga gcc cga ggg        748
95 Gln Arg Ala Val Ala Met Ser Ser Gly Asp Val Pro Gly Ala Arg Gly
96      190                 195                 200
98 gcc gtg gag ggg atc ctc atc cag cag gtg ttt gag tca ggc ggc tcc        796
99 Ala Val Glu Gly Ile Leu Ile Gln Gln Val Phe Glu Ser Gly Gly Ser
100 205                210                 215                 220
102 aag aag tgc atc cag gtt ggc ggg gag ttc tac act ccc agc aag ttc       844
103 Lys Lys Cys Ile Gln Val Gly Gly Glu Phe Tyr Thr Pro Ser Lys Phe
104                 225                 230                 235
106 gaa gac tcc ggc agt ggg aag aac aag gcc cgc agc agc agt ggc ccg       892
107 Glu Asp Ser Gly Ser Gly Lys Asn Lys Ala Arg Ser Ser Ser Gly Pro
108                 240                 245                 250
110 aag cct ctg gtt cga gcc aag gga gcc cag ggc gct gcc ccc ggt gga       940
111 Lys Pro Leu Val Arg Ala Lys Gly Ala Gln Gly Ala Ala Pro Gly Gly
112         255                 260                 265
114 ggt gag gct agg ctg ggc cag cag ggc agc gtt ccc gcc cct ctg gcc       988
115 Gly Glu Ala Arg Leu Gly Gln Gln Gly Ser Val Pro Ala Pro Leu Ala
116      270                 275                 280
118 ctc ccc agt gac ccc cag ctc cac cag aag aat gag gac gag tgt gcc      1036
119 Leu Pro Ser Asp Pro Gln Leu His Gln Lys Asn Glu Asp Glu Cys Ala
120 285                 290                 295                 300
122 gtg tgt cgg gac ggc ggg gag ctc atc tgc tgt gac ggc tgc cct cgg      1084
123 Val Cys Arg Asp Gly Gly Glu Leu Ile Cys Cys Asp Gly Cys Pro Arg
124                 305                 310                 315
126 gcc ttc cac ctg gcc tgc ctg tcc cct ccg ctc cgg gag atc ccc agt      1132
127 Ala Phe His Leu Ala Cys Leu Ser Pro Pro Leu Arg Glu Ile Pro Ser
128                 320                 325                 330
```

```
130 ggg acc tgg agg tgc tcc agc tgc ctg cag gca aca gtc cag gag gtg       1180
131 Gly Thr Trp Arg Cys Ser Ser Cys Leu Gln Ala Thr Val Gln Glu Val
132         335                 340                 345
134 cag ccc cgg gca gag gag ccc cgg ccc cag gag cca ccc gtg gag acc       1228
135 Gln Pro Arg Ala Glu Glu Pro Arg Pro Gln Glu Pro Pro Val Glu Thr
136         350                 355                 360
138 ccg ctc ccc ccg ggg ctt agg tcg gcg gga gag gag gta aga ggt cca       1276
139 Pro Leu Pro Pro Gly Leu Arg Ser Ala Gly Glu Glu Val Arg Gly Pro
140 365                 370                 375                 380
142 cct ggg gaa ccc cta gcc ggc atg gac acg act ctt gtc tac aag cac       1324
143 Pro Gly Glu Pro Leu Ala Gly Met Asp Thr Thr Leu Val Tyr Lys His
144                 385                 390                 395
146 ctg ccg gct ccg cct tct gca gcc ccg ctg cca ggg ctg gac tcc tcg       1372
147 Leu Pro Ala Pro Pro Ser Ala Ala Pro Leu Pro Gly Leu Asp Ser Ser
148         400                 405                 410
150 gcc ctg cac ccc cta ctg tgt gtg ggt cct gag ggt cag cag aac ctg       1420
151 Ala Leu His Pro Leu Leu Cys Val Gly Pro Glu Gly Gln Gln Asn Leu
152         415                 420                 425
154 gct cct ggt gcg cgt tgc ggg gtg tgc gga gat ggt acg gac gtg ctg       1468
155 Ala Pro Gly Ala Arg Cys Gly Val Cys Gly Asp Gly Thr Asp Val Leu
156         430                 435                 440
158 cgg tgt act cac tgc gcc gct gcc ttc cac tgg cgc tgc cac ttc cca       1516
159 Arg Cys Thr His Cys Ala Ala Ala Phe His Trp Arg Cys His Phe Pro
160 445                 450                 455                 460
162 gcc ggc acc tcc cgg ccc ggg acg ggc ctg cgc tgc aga tcc tgc tca       1564
163 Ala Gly Thr Ser Arg Pro Gly Thr Gly Leu Arg Cys Arg Ser Cys Ser
164                 465                 470                 475
166 gga gac gtg acc cca gcc cct gtg gag ggg gtg ctg gcc ccc agc ccc       1612
167 Gly Asp Val Thr Pro Ala Pro Val Glu Gly Val Leu Ala Pro Ser Pro
168                 480                 485                 490
170 gcc cgc ctg gcc cct ggg cct gcc aag gat gac act gcc agt cac gag       1660
171 Ala Arg Leu Ala Pro Gly Pro Ala Lys Asp Asp Thr Ala Ser His Glu
172         495                 500                 505
174 ccc gct ctg cac agg gat gac ctg gag tcc ctt ctg agc gag cac acc       1708
175 Pro Ala Leu His Arg Asp Asp Leu Glu Ser Leu Leu Ser Glu His Thr
176         510                 515                 520
178 ttc gat ggc atc ctg cag tgg gcc atc cag agc atg gcc cgt ccg gcg       1756
179 Phe Asp Gly Ile Leu Gln Trp Ala Ile Gln Ser Met Ala Arg Pro Ala
180 525                 530                 535                 540
182 gcc ccc ttc ccc tcc tgaccccaga tggccgggac atgcagctct gatgagagag       1811
183 Ala Pro Phe Pro Ser
184                 545
186 tgctgagaag gacacctcct tcctcagtcc tggaagccgg ccggctggga tcaagaaggg    1871
188 gacagcgcca cctcttgtca gtgctcggct gtaaacagct ctgtgtttct ggggacacca    1931
190 gccatcatgt gcctggaaat taaaccctgc cccacttctc tactctggaa gtccccggga    1991
192 gcctctcctt gcctggtgac ctactaaaaa tataaaaatt agctg                    2036
195 <210> SEQ ID NO: 2
196 <211> LENGTH: 545
197 <212> TYPE: PRT
```

```
198 <213> ORGANISM: HOMO SAPIENS
200 <400> SEQUENCE: 2
202 Met Ala Thr Asp Ala Ala Leu Arg Arg Leu Leu Arg Leu His Arg Thr
203 1               5                   10                  15
206 Glu Ile Ala Val Ala Val Asp Ser Ala Phe Pro Leu Leu His Ala Leu
207             20                  25                  30
210 Ala Asp His Asp Val Val Pro Glu Asp Lys Phe Gln Glu Thr Leu His
211         35                  40                  45
214 Leu Lys Glu Lys Glu Gly Cys Pro Gln Ala Phe His Ala Leu Leu Ser
215     50                  55                  60
218 Trp Leu Leu Thr Gln Asp Ser Thr Ala Ile Leu Asp Phe Trp Arg Val
219 65                  70                  75                  80
222 Leu Phe Lys Asp Tyr Asn Leu Glu Arg Tyr Gly Arg Leu Gln Pro Ile
223             85                  90                  95
226 Leu Asp Ser Phe Pro Lys Asp Val Asp Leu Ser Gln Pro Arg Lys Gly
227             100                 105                 110
230 Arg Lys Pro Pro Ala Val Pro Lys Ala Leu Val Pro Pro Pro Arg Leu
231         115                 120                 125
234 Pro Thr Lys Arg Lys Ala Ser Glu Glu Ala Arg Ala Ala Ala Pro Ala
235     130                 135                 140
238 Ala Leu Thr Pro Arg Gly Thr Ala Ser Pro Gly Ser Gln Leu Lys Ala
239 145                 150                 155                 160
242 Lys Pro Pro Lys Lys Pro Glu Ser Ser Ala Glu Gln Gln Arg Leu Pro
243             165                 170                 175
246 Leu Gly Asn Gly Ile Gln Thr Met Ser Ala Ser Val Gln Arg Ala Val
247             180                 185                 190
250 Ala Met Ser Ser Gly Asp Val Pro Gly Ala Arg Gly Ala Val Glu Gly
251         195                 200                 205
254 Ile Leu Ile Gln Gln Val Phe Glu Ser Gly Gly Ser Lys Lys Cys Ile
255     210                 215                 220
258 Gln Val Gly Gly Glu Phe Tyr Thr Pro Ser Lys Phe Glu Asp Ser Gly
259 225                 230                 235                 240
262 Ser Gly Lys Asn Lys Ala Arg Ser Ser Ser Gly Pro Lys Pro Leu Val
263             245                 250                 255
266 Arg Ala Lys Gly Ala Gln Gly Ala Ala Pro Gly Gly Gly Glu Ala Arg
267             260                 265                 270
270 Leu Gly Gln Gln Gly Ser Val Pro Ala Pro Leu Ala Leu Pro Ser Asp
271         275                 280                 285
274 Pro Gln Leu His Gln Lys Asn Glu Asp Glu Cys Ala Val Cys Arg Asp
275     290                 295                 300
278 Gly Gly Glu Leu Ile Cys Cys Asp Gly Cys Pro Arg Ala Phe His Leu
279 305                 310                 315                 320
282 Ala Cys Leu Ser Pro Pro Leu Arg Glu Ile Pro Ser Gly Thr Trp Arg
283             325                 330                 335
286 Cys Ser Ser Cys Leu Gln Ala Thr Val Gln Glu Val Gln Pro Arg Ala
287             340                 345                 350
290 Glu Glu Pro Arg Pro Gln Glu Pro Pro Val Glu Thr Pro Leu Pro Pro
291         355                 360                 365
294 Gly Leu Arg Ser Ala Gly Glu Glu Val Arg Gly Pro Pro Gly Glu Pro
```

```
295     370                     375                 380
298 Leu Ala Gly Met Asp Thr Thr Leu Val Tyr Lys His Leu Pro Ala Pro
299 385                     390                 395                 400
302 Pro Ser Ala Ala Pro Leu Pro Gly Leu Asp Ser Ser Ala Leu His Pro
303                     405                 410                 415
306 Leu Leu Cys Val Gly Pro Glu Gly Gln Gln Asn Leu Ala Pro Gly Ala
307                 420                 425                 430
310 Arg Cys Gly Val Cys Gly Asp Gly Thr Asp Val Leu Arg Cys Thr His
311         435                 440                 445
314 Cys Ala Ala Ala Phe His Trp Arg Cys His Phe Pro Ala Gly Thr Ser
315     450                 455                 460
318 Arg Pro Gly Thr Gly Leu Arg Cys Arg Ser Cys Ser Gly Asp Val Thr
319 465                 470                 475                 480
322 Pro Ala Pro Val Glu Gly Val Leu Ala Pro Ser Pro Ala Arg Leu Ala
323                 485                 490                 495
326 Pro Gly Pro Ala Lys Asp Asp Thr Ala Ser His Glu Pro Ala Leu His
327             500                 505                 510
330 Arg Asp Asp Leu Glu Ser Leu Leu Ser Glu His Thr Phe Asp Gly Ile
331         515                 520                 525
334 Leu Gln Trp Ala Ile Gln Ser Met Ala Arg Pro Ala Ala Pro Phe Pro
335     530                 535                 540
338 Ser
339 545
342 <210> SEQ ID NO: 3
343 <211> LENGTH: 1545
344 <212> TYPE: DNA
345 <213> ORGANISM: HOMO SAPIENS
348 <220> FEATURE:
349 <221> NAME/KEY: CDS
350 <222> LOCATION: (237)..(1283)
353 <220> FEATURE:
```

W--> 354 `<221> NAME/KEY: mat peptide`

```
355 <222> LOCATION: (1)..(348)
356 <223> OTHER INFORMATION: /product="AIR-2"
359 <400> SEQUENCE: 3
360 agagaaagtg aggtcttctc aggctcttaa gagcatggcg tttggtccag gctgtacccg     60
362 ctgctctcag ctgggcccgt gggtgggccg ggcgcccctg ctatagccag gaggtcaagg    120
364 atccactggg aatgccatgc tcatctttcg tccccagcat ggtttcttaa tggggtagaa    180
366 gcaggtcggg agagacctcc ctgggcctgg ccccactgcc ctgtgaggaa gggttc atg    239
367                                                             Met
368                                                             1
370 tgg ttg gtg tac agt tcc ggg gcc cct gga acg cag cag cct gca aga    287
371 Trp Leu Val Tyr Ser Ser Gly Ala Pro Gly Thr Gln Gln Pro Ala Arg
372             5                   10                  15
374 aac cgg gtt ttc ttc cca ata ggg atg gcc ccg ggg ggt gtc tgt tcg    335
375 Asn Arg Val Phe Phe Pro Ile Gly Met Ala Pro Gly Gly Val Cys Ser
376         20                  25                  30
378 aga cca gat gga tgg gga aca ggt ggt cag ggc aga att tca ggc cct    383
379 Arg Pro Asp Gly Trp Gly Thr Gly Gly Gln Gly Arg Ile Ser Gly Pro
```

RAW SEQUENCE LISTING ERROR SUMMARY   DATE: 10/05/2005
PATENT APPLICATION: **US/09/508,658D**   TIME: 08:43:06

Input Set : **A:\u012653-9SEQUENCE LISTING.txt**
Output Set: **N:\CRF4\10052005\I508658D.raw**

Invalid <213> Response:
Use of "Artificial" only as "<213> Organism" response is incomplete,
per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31
Seq#:32,33,34,35,36

VERIFICATION SUMMARY   DATE: 10/05/2005
PATENT APPLICATION: **US/09/508,658D**   TIME: 08:43:06

Input Set : **A:\u012653-9SEQUENCE LISTING.txt**
Output Set: **N:\CRF4\10052005\I508658D.raw**

L:36 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:1
L:354 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:3
L:577 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:5